Optimized and Scalable Climate Data Services (T258A):
International Research Institute for Climate and Society (IRI) Climate Data Library

M. Benno Blumenthal, M. Bell, R. Cousin, J. del Corral, H. Liu (jdcorral@iri.columbia.edu)

eLearning

- Web classroom references
- Tutorials
- Standalone platform

Faceted Search

- GCMD Keywords
- SWEET Classes
- IRI Terms
- CF Standard Names

Federation

- OPeNDAP data transfer
- THREDDS data catalog

Graphics

- PDF, PS, GIF, JPG
- geoTIFF
- Animations

Data Download/Upload

- ASCII, binary, netCDF, tables, OPeNDAP, GIS (Download)
- Tables, GIS (Upload)

Data Center

- squid caching
- caching server(s)
- redundant compute servers
- redundant data bases
- load balancing

Spatio-Temporal Analysis

- Regridding
- Time Series Analysis
- Spatial Averaging
- Gridding of point data

Map Services

- OGC WMS
- KML (GoogleEarth)
- ArcGIS
- SERVIR

User Platform

- mirror config.
- standalone config.
- laptop
- bootable USB disk

Statistical Analysis

- EOFs
- Correlations
- Clustering
- Singular Value Decomposition

Connection to Applications

- OPeNDAP clients
- Matlab
- GrADS

Tailored Maprooms

- Public Health
- Disaster
- Regional

Programming Interface

- Custom procedures
- Debugging

Data Integration

- Observations
- Models
- Public Health
- Agriculture
- Disasters

Data Repository

- Stations, Buoys, Satellites (1D, 2D, 3D, 4D)
- Model Forecasts/Hindcasts (nD)
- Demographics
- OGC GIS features
- Agriculture production
- Public Health records

Introduction:

IRI Climate Data Library (http://iridl.ldeo.columbia.edu) is designed to optimize the display, analysis, and retrieval of climate datasets. These datasets range from simple station observations, to multi-ensemble climate model results, to high-resolution satellite measurements, to GIS representations of geographic entities. These datasets are represented in a consistent multi-dimensional framework. As a result, station observations can easily be compared with climate model results, and satellite measurements. Gridded data can be spatially averaged over discrete geographic entities. The Climate Data Library is accessible with a browser connected to the internet or local area network. The data selection, processing, and analysis are performed by the Climate Data Library servers. The resulting images or data files are sent back to the client's desktop. This model optimizes the use of internet bandwidth. The software required is open source and available in most parts of the world.

Summary:

Scalable climate data services can be based on the use of common data standards. While we don't expect everyone to use our software system, we do advocate the use of many of the data and metadata standards that we incorporate in our system. We incorporate caching and threaded interpreted instructions to optimize access, display, and retrieval of data. These can serve as models for the climate community.